



Europäisches  
Patentamt

European  
Patent Office

Office européen  
des brevets

MAILED 19 APR 2004

WIPO

PCT

Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-  
gen stimmen mit der  
ursprünglich eingereichten  
Fassung der auf dem näch-  
sten Blatt bezeichneten  
europäischen Patentanmel-  
dung überein.

The attached documents  
are exact copies of the  
European patent application  
described on the following  
page, as originally filed.

Les documents fixés à  
cette attestation sont  
conformes à la version  
initialement déposée de  
la demande de brevet  
européen spécifiée à la  
page suivante.

BEST AVAILABLE COPY

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03100947.5

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

R C van Dijk



Anmeldung Nr:  
Application no.: 03100947.5  
Demande no:

Anmeldetag:  
Date of filing: 08.04.03  
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Philips Intellectual Property & Standards  
GmbH  
Steindamm 94  
20099 Hamburg  
ALLEMAGNE  
Koninklijke Philips Electronics N.V.  
Groenewoudseweg 1  
5621 BA Eindhoven  
PAYS-BAS

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:  
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.  
If no title is shown please refer to the description.  
Si aucun titre n'est indiqué se referer à la description.)

Verfahren und Vorrichtung zur Reduktion eines Störgeräuschsignalanteils in einem  
Mikrofonsignal

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s)  
revendiquée(s)  
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/  
Classification internationale des brevets:

H04R3/00

Am Anmeldetag benannte Vertragsstaaten/Contracting states designated at date of  
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL  
PT SE SI SK TR LI

## BESCHREIBUNG

Verfahren und Vorrichtung zur Reduktion eines Störgeräuschsignalanteils in einem Mikrofonsignal

Die Erfindung betrifft ein Verfahren zur Reduktion eines Störgeräuschsignalanteils in einem  
5 Mikrofonsignal. Außerdem betrifft die Erfindung noch eine Vorrichtung zur Reduktion eines Störgeräuschsignalanteils in einem Mikrofonsignal.

Derartige Verfahren sind insbesondere für die Qualitätsverbesserung von Sprachsignalen, die einer Spracherkennungseinrichtung oder einer Telekommunikationseinrichtung zugeführt  
10 werden von großer Bedeutung. Ein wichtiges Anwendungsbeispiel aus dem Telekommunikationsbereich sind Freisprecheinrichtungen, deren Verwendung zum Telefonieren in Automobilen heutzutage sogar vom Gesetzgeber vorgeschrieben wird. Mit Hilfe solcher Freisprecheinrichtungen ist es dem Fahrer möglich mit einem entfernten Gesprächsteilnehmer zu kommunizieren, ohne dabei seine Hände vom Steuer nehmen zu müssen und damit seine  
15 Aufmerksamkeit vom Straßenverkehr abzuwenden.

An dem Beispiel der Freisprecheinrichtung lassen sich gut die beiden Arten von Störgeräuschen erläutern, zwischen denen man prinzipiell unterscheidet, und deren Beseitigung aus dem zum entfernten Gesprächsteilnehmer übertragenen Sprachsignal die Aufgabe der hier  
20 besprochenen Verfahren ist.

Zum einen gibt es die Störgeräusche, die von einer oder mehreren bekannten Schallquellen stammen. Im Falle der Freisprecheinrichtungen im Auto sind dies zum Beispiel die Geräusche, die von dem Lautsprecher der Freisprecheinrichtung oder den Lautsprechern einer Audio  
25 Anlage erzeugt werden. Gelangt beispielsweise das von dem Lautsprecher der Freisprecheinrichtung erzeugte Sprachsignal des entfernten Gesprächsteilnehmers zum Mikrofon und wird es nicht aus dem Mikrofonsignal entfernt, so vernimmt der entfernte Gesprächsteilnehmer ein Echo seiner eigenen Stimme, was als sehr unangenehm wahrgenommen wird. Die Ver-

fahren, die dazu verwendet werden derartige Störgeräuschanteile aus dem Mikrofonsignal zu entfernen, erfordern die Kenntnis über das Signal, welches das Störgeräusch erzeugt. In dem oben beschriebenen Beispiel ist dies das Sprachsignal des entfernten Gesprächsteilnehmers, welches dem Lautsprecher der Freisprecheinrichtung zugeführt wird. Solche Verfahren sind  
5 beispielsweise in der EP 0 948 237 A2 und in der DE 41 06 405 A1 beschrieben.

Zu der zweiten Art von Störgeräuschen zählt man solche, über deren Erzeugung man keine genaue Kenntnis hat und die in der Regel von einer Vielzahl nicht genau definierter Schallquellen erzeugt werden. Typische Umgebungsgeräusche zählen zu dieser Art von Störgeräuschen. Betrachtet man wiederum das Beispiel der Freisprecheinrichtung im Auto, so zählt  
10 man das Fahrgeräusch des Autos zu dieser Art von Störgeräuschen. Eine große Gruppe von Verfahren zur Reduktion von Störgeräuschen dieser Art basiert auf der Schätzung des Störgeräuschanteils auf der Grundlage des Mikrofonsignals. Mit Hilfe dieser Schätzung wird, beispielsweise mittels der Methode der spektralen Subtraktion, der Störgeräuschsignalanteil in  
15 dem Mikrofonsignal reduziert. Ein Verfahren aus dieser Gruppe ist beispielsweise in der US 6,363,345 B1 beschrieben. Bei der Schätzung des Störgeräuschanteils aus dem Mikrofonsignal stellt sich jedoch das Problem, dass innerhalb des Mikrofonsignals solche Signalabschnitte detektiert werden müssen, in denen nur ein Störgeräuschsignalanteil und kein Nutzsignalanteil vorliegt. Im Falle der Freisprecheinrichtung im Auto wären dies solche  
20 Signalabschnitte im Mikrofonsignal, die keinen Sprachsignalanteil enthalten. Sofern aufgrund des Szenarios derartige Signalabschnitte überhaupt präsent sind, ist zur Detektion dieser Signalabschnitte ein zusätzlicher Signalverarbeitungsschritt, die sogenannte Voice-Activity-Detection (VAD) notwendig. Insbesondere in dem Fall eines schlechten Störgeräuschabstandes (Signal-to-Noise-Ratio, SNR) in dem Mikrofonsignal liefert die VAD jedoch häufig  
25 nur unzuverlässige Ergebnisse. Darüber hinaus muss die Annahme gemacht werden, dass die in dem sprachsignalfreien Abschnitt gemachte Störgeräuschsignalabschätzung auch zu späteren Zeitpunkten gültig ist. Insbesondere bei sich zeitlich schnell ändernden Störgeräuschen in Kombination mit langen Sprachsignalabschnitten ist diese Annahme jedoch nur eine unzureichende Näherung.

Aufgabe der vorliegenden Erfindung ist es daher, ein Verfahren zur Reduktion eines Störgeräuschsignalanteils in einem Mikrofonsignal anzugeben, welches mit einem geringen Signalverarbeitungsaufwand eine gute Schätzung des Störgeräuschsignalanteils und damit eine gute

5 Reduktion des Störgeräuschsignalanteils in dem Mikrofonsignal ermöglicht.

Die oben formulierte Aufgabe wird erfindungsgemäß durch ein Verfahren mit den Schritten nach Anspruch 1 gelöst. Die abhängigen Unteransprüche enthalten vorteilhafte Ausgestaltungen und Weiterbildungen des Verfahrens gemäß Anspruch 1.

10

Gemäß des erfindungsgemäßen Verfahrens erfolgt die Bestimmung des Störgeräuschreferenzsignals oder der Störgeräuschreferenzsignale, welche als Basis für die Schätzung des Störgeräuschsignalanteils in dem interessierenden Mikrofonsignal dienen, mittels jeweils eines

invers, also als Mikrofon, betriebenen Lautsprechers.

15

Dabei ist der Lautsprecher so geeignet positioniert, dass der von der Störgeräuschquelle stammende Signalanteil in dem zugehörigen Störgeräuschreferenzsignal mindestens so hoch ist wie der von der Sprachsignalquelle stammende Signalanteil. Verwendet man das in der Signalverarbeitung gängige Maß SNR und identifiziert man den von der Sprachsignalquelle

20 stammenden Signalanteil in diesem Sinne als Signal und den von der Störgeräuschquelle stammenden Signalanteil als Rauschen, so entspricht dies einem SNR kleiner gleich Null.

Vorzugsweise ist der von der Störgeräuschquelle stammende Signalanteil in dem zugehörigen Störgeräuschreferenzsignal sogar doppelt so hoch wie der von der Sprachsignalquelle stammende Signalanteil, was einem SNR von etwa  $-6$  entspricht. Durch eine derartige

25

Positionierung der Lautsprecher ist die Information über den Störgeräuschsignalanteil, die man aus den Lautsprechersignalen gewinnen kann, nur wenig durch Sprachsignalanteile verfälscht. Insbesondere in Situationen, wo bereits ein oder mehrere Lautsprecher als Komponente eines Audio Systems vorhanden sind, entfällt bei dem erfindungsgemäßen Verfahren die Notwendigkeit der Installation zusätzlicher Mikrofone.

- Die Ermittlung der Schätzung des Störgeräuschsignalanteils aus den Lautsprechersignalen, welche auch als Störgeräuschreferenzsignale bezeichnet werden, erfolgt in Abhängigkeit davon, ob nur ein oder mehrere solche Signale vorliegen, in einem bzw. in zwei Schritten. Bei nur einem zur Verfügung stehenden Störgeräuschreferenzsignal wird ein Verfahren der Signal-
- 5 Schätzungs-Theorie, beispielsweise eine rekursive Noise-Schätzung, auf dieses Signal angewendet und damit direkt die Schätzung des Störgeräuschsignalanteils ermittelt. Im Falle von mehr als einem Störgeräuschreferenzsignal wird im ersten Schritt auf jedes dieser Signale ein Verfahren der Signal-Schätzungs-Theorie, beispielsweise die rekursive Noise-Schätzung angewendet und damit jeweils eine vorläufige Schätzung des Störgeräuschsignalanteils
- 10 ermittelt. Im zweiten Schritt werden diese vorläufigen Schätzungen des Störgeräuschsignalanteils dann durch lineare Überlagerung kombiniert, wodurch man schließlich zu der gesuchten Schätzung des Störgeräuschsignalanteils gelangt. Die lineare Überlagerung wird vorzugsweise so durchgeführt, dass zunächst die vorläufigen Schätzungen des Störgeräuschsignalanteils mit jeweils einem Gewichtungsfaktor multipliziert und anschließend die so gewon-
- 15 nenen gewichteten vorläufigen Schätzungen des Störgeräuschsignalanteils aufsummiert werden. Die Gewichtungsfaktoren spiegeln dabei die Übertragungskanal-Charakteristik des entsprechenden Lautsprechersignals wider. Es lässt sich qualitativ die Aussage treffen, je weiter entfernt der Lautsprecher von der Sprachsignalquelle positioniert ist, desto größer ist die Dämpfung des Sprachsignals in diesem Lautsprecher und desto größer ist folglich der zugehörige
- 20 Gewichtungsfaktor.

- Nachdem die Schätzung des Störgeräuschsignalanteils ermittelt wurde, wird diese von dem Mikrofonsignal, beispielsweise mittels Optimalfilterung abgezogen, wodurch man schließlich das bereinigte, also das um den Störgeräuschsignalanteil reduzierte, Mikrofonsignal erhält. Bei
- 25 dem Verfahren der Optimalfilterung wird auf der Grundlage der Schätzung des Störgeräuschsignalanteils und des Mikrofonsignals der Frequenzgang eines Filters, des sogenannten Optimalfilters oder Wienerfilters, berechnet und der Abzug des Störgeräuschsignalanteils von dem Mikrofonsignal durch Anwendung dieses Filters auf das Mikrofonsignal durchgeführt. Dies kann sowohl im Zeitbereich als auch im Frequenzbereich erfolgen. Weitere Verfahren zum

Abzug des Störgeräuschsignalanteils von dem Mikrofonsignal sind beispielsweise die spektrale Subtraktion und die nichtlineare spektrale Subtraktion.

5 In einer anderen Ausgestaltung des erfindungsgemäßen Verfahrens wird zusätzlich zu den mit den Lautsprechern aufgenommenen Störgeräuschreferenzsignalen und der daraus ermittelten Schätzung des Störgeräuschsignalanteils, welche im folgenden als erste Schätzung bezeichnet wird, auch das Mikrofonsignal selber zur Ermittlung einer zweiten Schätzung des Störgeräuschsignalanteils herangezogen. In einem weiteren Schritt werden die erste und die zweite Schätzung dann, ganz analog den vorläufigen Schätzungen bei mehreren Störgeräuschreferenz-  
10 signalen, durch lineare Überlagerung kombiniert und somit die gesuchte Schätzung des Störgeräuschsignalanteils ermittelt.

Für das mit Hilfe des erfindungsgemäßen Verfahrens erhaltene bereinigte Mikrofonsignal sind die verschiedensten Verwendungen denkbar. So kann es beispielsweise einer Telekommuni-  
15 kationseinrichtung zugeführt und damit einem entfernten Gesprächspartner übermittelt werden, wodurch für diesen die Qualität des empfangenen Sprachsignals erhöht wird. In einer weiteren Verwendung kann das bereinigte Mikrofonsignal einem Spracherkennungssystem zugeführt werden, wodurch die Erkennungsleistung dieses Systems erhöht wird.

20 In einer weiteren Ausgestaltung des erfindungsgemäßen Verfahrens erfolgt die Aufnahme des Mikrofonsignals und des mindestens einen Störgeräuschreferenzsignals in einem Verkehrsmittel, beispielsweise einem Kraftfahrzeug, und die dabei verwendeten Lautsprecher sind Teil eines bereits vorhandenen Lautsprechersystems. Insbesondere in einem Kraftfahrzeug ist dies besonders vorteilhaft, da die Lautsprecher dort in der Regel schon so positioniert sind, dass  
25 der Störgeräuschsignalanteil in dem mit ihnen aufgenommenen Signal mindestens so hoch ist, wie der von einem auf dem Fahrersitz sitzenden Sprecher stammende Sprachsignalanteil.

Weitertin betrifft die Erfindung eine Vorrichtung zur Ausführung des Verfahrens nach Anspruch 1. Die Vorrichtung umfasst einen Signalprozessor, auf welchem die Bestimmung der

Schätzung des Störgeräuschsignalanteils und der Abzug dieser Schätzung von dem Mikrofon-signal durchgeführt wird. Weiterhin umfasst die Vorrichtung mindestens ein Mikrofon, welches mit dem Signalprozessor gekoppelt ist. Diese Kopplung kann beispielsweise mittels einer Leitung oder drahtlos vorgenommen werden und es ist in der Regel ein sogenannter Codec  
5 zur Analog-Digital-Wandlung des Mikrofonsignals zwischengeschaltet. Ebenso umfasst die Vorrichtung mindestens einen Lautsprecher, welcher als Mikrofon betrieben wird und welcher ebenfalls mit dem Signalprozessor gekoppelt ist. Auch dabei kann die Kopplung beispielsweise mittels einer Leitung oder drahtlos erfolgen und es kann ein Codec zur Analog-Digital-Wandlung des Lautsprechersignals zwischengeschaltet werden. Auf dem Signalpro-  
10 zessor können neben den zum erfindungsgemäßen Verfahren gehörenden Verarbeitungsschritte auch noch weitere Datenverarbeitungsschritte ausgeführt werden. Der Signalprozessor kann insbesondere auch Teil einer bereits vorhandenen Datenverarbeitungseinrichtung sein und zusätzlich für das erfindungsgemäße Verfahren verwendet werden.

15 Die Erfindung ist nachfolgend anhand eines bevorzugten Ausführungsbeispiels unter Bezugnahme auf die Abbildungen noch eingehend veranschaulicht. Dabei zeigen:

Fig. 1 ein Blockschaltbild zur Darstellung des erfindungsgemäßen Verfahrens,

20 Fig. 2 ein Flussdiagramm, welches die Ermittlung einer vorläufigen Schätzungen eines Störgeräuschsignalanteils veranschaulicht,

Fig. 3 ein Flussdiagramm, welches die Kombination der vorläufigen Schätzungen des Störgeräuschsignalanteils zur Ermittlung einer Schätzung des Störgeräuschsignalanteils  
25 veranschaulicht und

Fig. 4 ein Flussdiagramm, welches den Abzug der Schätzung des Störgeräuschsignalanteils von einem Mikrofonsignal veranschaulicht.



Figur 1 zeigt ein Blockschaltbild einer Anordnung zur Durchführung des erfindungsgemäßen Verfahrens. Ein Mikrofonsignal  $x$ , welches mittels des erfindungsgemäßen Verfahrens von einem Störgeräuschsignalanteil befreit werden soll, wird mit Hilfe eines Mikrofons 101 aufgezeichnet und einer Abzugseinheit 501, welche den Abzug der Schätzung des Störgeräuschsignalanteils von dem Mikrofonsignal durchführt, zugeführt. Lautsprecher 201, 202 und 203 werden in bekannter Weise als Mikrofone verwendet, und mit ihnen werden Störgeräuschreferenzsignale  $x_1$ ,  $x_2$  und  $x_3$  aufgezeichnet. Dabei ist die beispielhafte Auswahl von drei Lautsprechern und dementsprechend drei Störgeräuschreferenzsignalen keineswegs zwingend. Vielmehr ist, ausgehend von mindestens einem Lautsprecher und dementsprechend einem Störgeräuschreferenzsignal, die Anzahl beliebig und höchstens durch den sich ergebenden Signalverarbeitungsaufwand begrenzt. Die drei Störgeräuschreferenzsignale  $x_1$ ,  $x_2$  und  $x_3$  werden nun jeweils einer Schätzungseinheit 301, 302 bzw. 303 zugeführt. In diesen Schätzungseinheiten wird jeweils eine vorläufige Schätzung des Störgeräuschsignalanteils ermittelt. Diese vorläufigen Schätzungen des Störgeräuschsignalanteils, welche in Fig. 1 mit  $N_1$ ,  $N_2$  und  $N_3$  bezeichnet sind, werden im folgenden einer Kombinationseinheit 401 zugeführt. Diese Kombinationseinheit 401 kombiniert die vorläufigen Schätzungen des Störgeräuschsignalanteils und ermittelt damit eine Schätzung des Störgeräuschsignalanteils, welche in Fig. 1 mit  $N$  bezeichnet ist. Diese Schätzung des Störgeräuschsignalanteils wird nun, neben dem Mikrofonsignal, als zweites Eingangssignal der Abzugseinheit 501 zugeführt. Innerhalb dieser Abzugseinheit 501 wird die Schätzung des Störgeräuschsignalanteils von dem Mikrofonsignal abgezogen und somit ein bereinigtes Signal  $x'$  ermittelt.

Figur 2 zeigt ein Flussdiagramm, welches die Funktionsweise der Schätzungseinheit 301 veranschaulicht. Innerhalb dieser Schätzungseinheit 301 wird die vorläufige Schätzung des Störgeräuschsignalanteils  $N_1$  aus dem, mittels des Lautsprechers 201 aufgenommenen, Signal  $x_1$  berechnet wird. Die Funktionsweise der Schätzungseinheiten 302 und 303 ist damit identisch. Zunächst wird das Signal  $x_1$  mittels einer Analog-Digital-Wandlung 310 mit einer Abtastrate von 8 kHz digitalisiert. Danach wird mittels einer sogenannten Rahmenbildung (framing) 311 ein Block von  $M$  digitalen Sample Werten des Signals  $x_1$  gebildet. Und zwar

setzt sich dieser Block aus den letzten M-B Sample Werten des vorhergehenden Blocks und aus den letzten B aktuellen Sample Werten des Signals  $x_1$  zusammen. Die Signalverarbeitung erfolgt also in aufeinanderfolgenden Blöcken mit M Sample Werten, die um M-B Sample Werte überlappen, wobei jeweils B aktuelle Sample Werte verarbeitet werden. Wählt man

5 M=256 und B=128, so entspricht, bei einer Abtastrate von 8 kHz, ein Block einer Zeitdauer von 32 ms und die aufeinanderfolgenden Blöcke überlappen um 16 ms, also um 50%. In einer nun folgenden Fensterung 312 werden die M Sample Werte des Blocks mit den Funktionswerten einer Fensterfunktion, beispielsweise einer Hammingfunktion, multipliziert, um bei dem anschließenden Übergang in den Frequenzbereich störende Einflüsse durch die Rahmen-

10 bildung zu vermindern. Die auf diese Weise ermittelten „gefensterten“ Sample Werte werden nun mittels einer diskreten Fouriertransformation 313 in den Frequenzbereich transformiert. In einem anschließenden Verarbeitungsschritt 314 wird das Absolutquadrat der M komplexen Fourierkoeffizienten gebildet, womit man das Leistungsspektrum  $P_1(f,i)$  erhält. Dabei ist f die Frequenz und i der Index des aktuellen Blocks, der über die Blocklänge und die Abtastrate

15 mit der Zeit zusammenhängt. Dieses Leistungsspektrum wird nun mittels einer rekursiven Glättung 315 gemäß der Formel

$$N_1(f,i) = \alpha \cdot N_1(f,i-1) + (1-\alpha) \cdot P_1(f,i)$$

geglättet, womit man die vorläufige Schätzung des Störgeräuschsignalanteils im Frequenzbereich  $N_1(f,i)$  erhält. Der Glättungs-Filter-Koeffizient  $\alpha$  ist dabei ein zu optimierender

20 Parameter des Verfahrens. Ein typischer Wert für  $\alpha$  beispielsweise der Wert 0,99. An dieser Stelle sei angemerkt, dass die Bestimmung der vorläufigen Schätzung des Störgeräuschsignalanteils nicht zwingend im Frequenzbereich erfolgen muss. Vielmehr sind auch Realisierungen im Zeitbereich denkbar.

25 Figur 3 zeigt ein Flussdiagramm zur Veranschaulichung der Funktionsweise der Kombinationseinheit 401. Die vorläufigen Schätzungen des Störgeräuschsignalanteils  $N_1$ ,  $N_2$  und  $N_3$ , die, auf die oben beschriebene Weise, in den Schätzungseinheiten 301, 302 und 303 ermittelt wurden, werden zunächst jeweils mit einem Gewichtungsfaktor  $\beta_1$ ,  $\beta_2$  bzw.  $\beta_3$  multipliziert. Diese Gewichtungsfaktoren sind wiederum zu optimierende Parameter des erfindungsgemäßen Ver-

fahrens, und sie spiegeln die Übertragungskanal-Charakteristik des entsprechenden Lautsprechersignals wider. Es lässt sich qualitativ die Aussage treffen, je weiter entfernt der Lautsprecher von der Sprachsignalquelle positioniert ist, desto größer ist die Dämpfung des Sprachsignals in diesem Lautsprecher und desto größer ist folglich der zugehörige Gewichtungsfaktor  $\beta$ . Nachdem alle vorläufigen Schätzungen des Störgeräuschsignalanteils mit ihren jeweiligen Gewichtungsfaktoren multipliziert wurden, ergibt sich die Schätzung des Störgeräuschsignalanteils  $N$  als Summe dieser Produkte:

$$N(f, i) = \sum_k \beta_k \cdot N_k(f, i)$$

Es sei angemerkt, dass im Falle von nur einem Lautsprecher und dementsprechend nur einem Störgeräuschreferenzsignal der Verarbeitungsschritt innerhalb der Schätzungseinheit 401 wegfällt und die vorläufige Schätzung des Störgeräuschsignalanteils  $N_1(f, i)$  mit der Schätzung des Störgeräuschsignalanteils  $N(f, i)$  identisch ist.

In Figur 4 ist die Funktionsweise der Abzugseinheit 501, in welcher der letzte Schritt des erfindungsgemäßen Verfahrens, der Abzug der Schätzung des Störgeräuschsignalanteils von dem Mikrofonsignal, durchgeführt wird, anhand eines Flussdiagramms veranschaulicht. Zunächst wird das Mikrofonsignal  $x$ , in Analogie zu dem Lautsprechersignal  $x_1$  in Figur 2, einer Analog-Digital-Wandlung 510, einer Rahmenbildung 511, einer Fensterung 512, einer Transformation in den Frequenzbereich 513, und einer Berechnung des Leistungsspektrums  $P(f, i)$  514 als Absolutquadrat der komplexen Fourierkoeffizienten, unterworfen. Zusätzlich zu dem Leistungsspektrum wird in einem Verarbeitungsschritt 515 nun auch noch die Phase  $\varphi(f, i)$  der komplexen Fourierkoeffizienten  $X$  berechnet. Aus der in der Kombinationseinheit 401 bestimmten Schätzung des Störgeräuschsignalanteils  $N(f, i)$  und dem Leistungsspektrum des Mikrofonsignals  $P(f, i)$  wird nun mittels einer Nichtlinearen-Spektralen-Subtraktion 516 gemäß der Formel

$$P'(f, i) = \max\{P(f, i) - a(f, i) \cdot N(f, i), b \cdot N(f, i)\}$$

ein bereinigtes Leistungsspektrum  $P'(f, i)$  berechnet. Dabei sind der sogenannte Overestimation-Faktor  $a(f, i)$  und der sogenannte Floor-Faktor  $b$  zu optimierende Parameter

des erfindungsgemäßen Verfahrens. Zur Methode der Nichtlinearen-Spektralen-Subtraktion sei auf die Referenz Bouquin, R.L., „Enhancement of noisy speech signals: Applications to mobile radio communications“, Speech Communication, vol. 18, 1996 verwiesen. Aus dem bereinigten Leistungsspektrum und der zuvor berechneten unveränderten Phase  $\varphi(f,i)$  wird nun

5 in dem Verarbeitungsschritt 517 gemäß der Gleichung

$$X'(f,i) = \sqrt{P'(f,i)} \cdot e^{i\varphi(f,i)}$$

ein bereinigtes Spektrum komplexer Fourierkoeffizienten  $X'(f,i)$  berechnet. Schließlich ergibt sich aus diesem bereinigten Spektrum nach einer inversen Fouriertransformation 518 und einer der Rahmenbildung inversen Prozedur 519 gemäß der sogenannten Overlap-Add-Methode

10 das bereinigte Mikrofonsignal  $x'$ . Auch an dieser Stelle sei nochmals angemerkt, dass nicht zwangsweise ein Subtraktionsverfahren im Frequenzbereich gewählt werden muss, sondern durchaus auch Verfahren im Zeitbereich denkbar sind.

PATENTANSPRÜCHE

1. Verfahren zur Reduktion eines Störgeräuschsignalanteils in einem Mikrofonsignal, welches den von wenigstens einer Störgeräuschquelle stammenden Störgeräuschsignalanteil und einen von einer Sprachsignalquelle stammenden Sprachsignalanteil enthält, folgende Schritte aufweisend:

- 5       • Aufnahme des, den Störgeräuschsignalanteil und den Sprachsignalanteil enthaltenden, Mikrofonsignals,
- Aufnahme mindestens eines Störgeräuschreferenzsignals mittels jeweils eines invers betriebenen Lautsprechers, wobei der oder die Lautsprecher so positioniert sind, dass der von den Störgeräuschquellen stammende Signalanteil in dem jeweiligen
- 10       Störgeräuschreferenzsignal mindestens so hoch ist, wie der von der Sprachsignalquelle stammende Signalanteil in diesem Störgeräuschreferenzsignal,
- Im Falle nur eines Störgeräuschreferenzsignals, Bestimmung einer Schätzung des Störgeräuschsignalanteils aus dem Störgeräuschreferenzsignal mittels einer Methode der Signal-Schätzungs-Theorie,
- 15       • Im Falle von mehr als einem Störgeräuschreferenzsignal, Bestimmung jeweils einer vorläufigen Schätzung des Störgeräuschsignalanteils aus jedem der Störgeräuschreferenzsignale mittels einer Methode der Signal-Schätzungs-Theorie und anschließend Bestimmung der Schätzung des Störgeräuschsignalanteils in dem Mikrofonsignal durch Kombination dieser vorläufigen Schätzungen des
- 20       Störgeräuschsignalanteils,
- Reduktion des Störgeräuschsignalanteils in dem Mikrofonsignal durch Abzug der Schätzung des Störgeräuschsignalanteils von dem Mikrofonsignal.

2. Verfahren nach Anspruch 1,

dadurch gekennzeichnet,

dass, in einem weiteren Verfahrensschritt zusätzlich zu der Bestimmung einer ersten Schätzung des Störgeräuschsignalanteils mittels mindestens eines Störgeräuschreferenzsignals, eine

- 5 Bestimmung einer zweiten Schätzung des Störgeräuschsignalanteils mittels des Mikrofonsignals selbst durchgeführt wird und aus einer Linearkombination der ersten und der zweiten Schätzung des Störgeräuschsignalanteils eine dritte Schätzung ermittelt wird, und dass die Reduktion des Störgeräuschsignalanteils in dem Mikrofonsignal durch Abzug dieser Schätzung von dem Mikrofonsignal erfolgt.

10

3. Verfahren nach Anspruch 1,

dadurch gekennzeichnet,

dass im Falle von mehr als einem Störgeräuschreferenzsignal die Kombination der vorläufigen Schätzungen des Störgeräuschsignalanteils aus der Multiplikation einer jeden vorläufigen

- 15 Schätzungen des Störgeräuschsignalanteils mit jeweils einem Gewichtungsfaktor und anschließender Summation der so erhaltenen gewichteten vorläufigen Schätzungen des Störgeräuschsignalanteils besteht.

4. Verfahren nach einem der Ansprüche 1 bis 3,

20 dadurch gekennzeichnet,

dass der Abzug der Schätzung des Störgeräuschsignalanteils von dem Mikrofonsignal mittels Optimalfilterung durchgeführt wird.

5. Verfahren nach einem der Ansprüche 1 bis 3,

25 dadurch gekennzeichnet,

dass der Abzug der Schätzung des Störgeräuschsignalanteils von dem Mikrofonsignal mittels der Methode der spektralen Subtraktion durchgeführt wird.

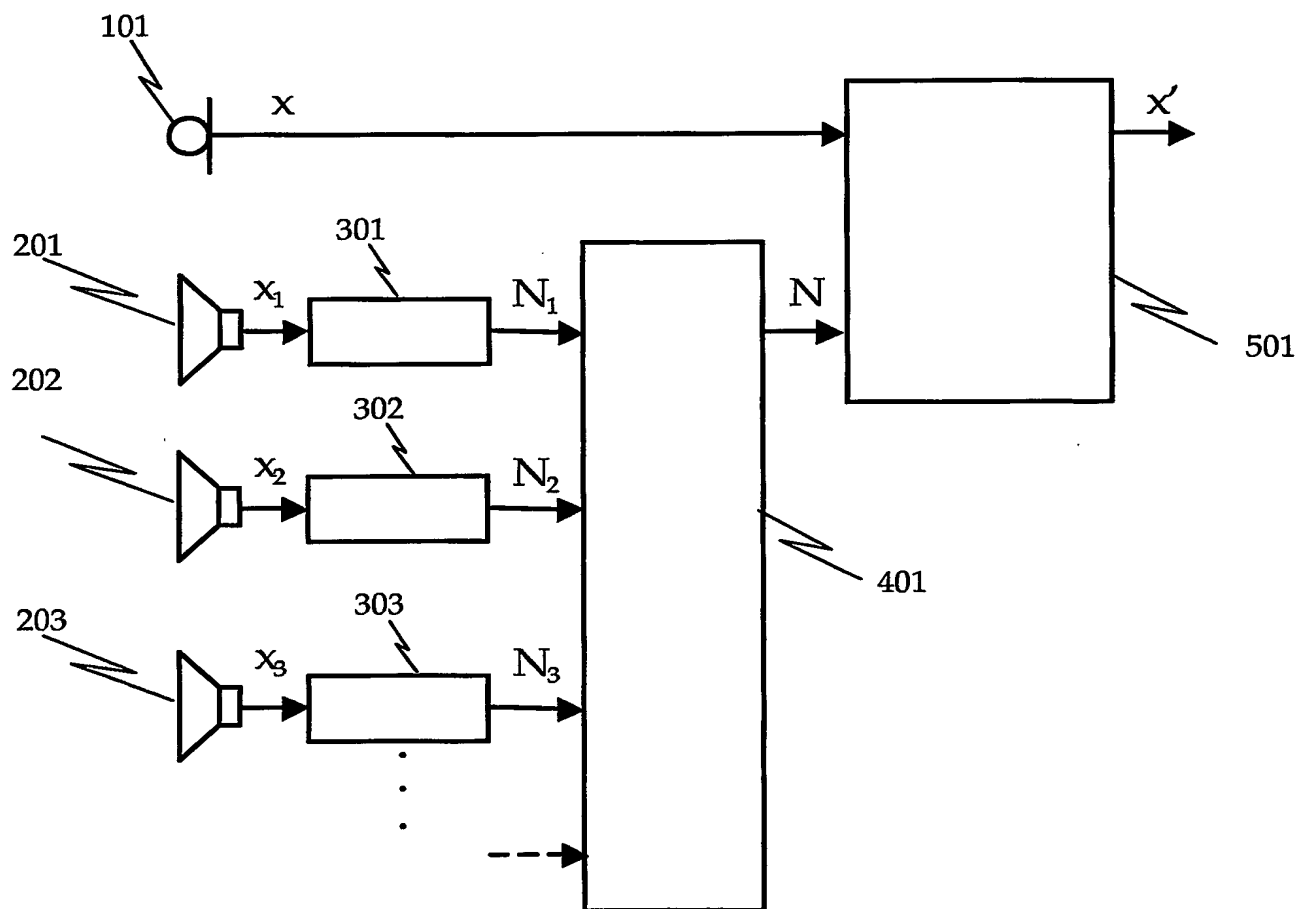
6. Verfahren nach einem der Ansprüche 1 bis 5,  
dadurch gekennzeichnet,  
dass das um den Störgeräuschsignalanteil reduzierte Mikrofonsignal einer  
5 Spracherkennungseinrichtung zugeführt wird.
7. Verfahren nach einem der Ansprüche 1 bis 5,  
dadurch gekennzeichnet,  
dass das um den Störgeräuschsignalanteil reduzierte Mikrofonsignal einer  
10 Telekommunikationseinrichtung zugeführt wird.
8. Verfahren nach einem der Ansprüche 1 bis 7,  
dadurch gekennzeichnet,  
dass die Aufnahme des Mikrofonsignals und des mindestens einen  
15 Störgeräuschreferenzsignals in einem Verkehrsmittel erfolgt und der oder die dabei  
verwendeten Lautsprecher Teil eines in dem Verkehrsmittel vorhandenen  
Lautsprechersystems sind.
9. Vorrichtung zur Ausführung des Verfahrens nach Anspruch 1, welche mindestens folgende  
20 Komponenten umfasst:
- Einen Signalprozessor zur Bestimmung der Schätzung des Störgeräuschsignalanteils  
und zum Abzug dieser Schätzung von dem Mikrofonsignal.
  - Mindestens ein Mikrofon, welches mit dem Signalprozessor gekoppelt und als  
Aufnahmegerät für das Mikrofonsignal vorgesehen ist.
  - 25 • Mindestens einen Lautsprecher, welcher mit dem Signalprozessor gekoppelt und als  
Aufnahmegerät für das Störgeräuschreferenzsignal vorgesehen ist.

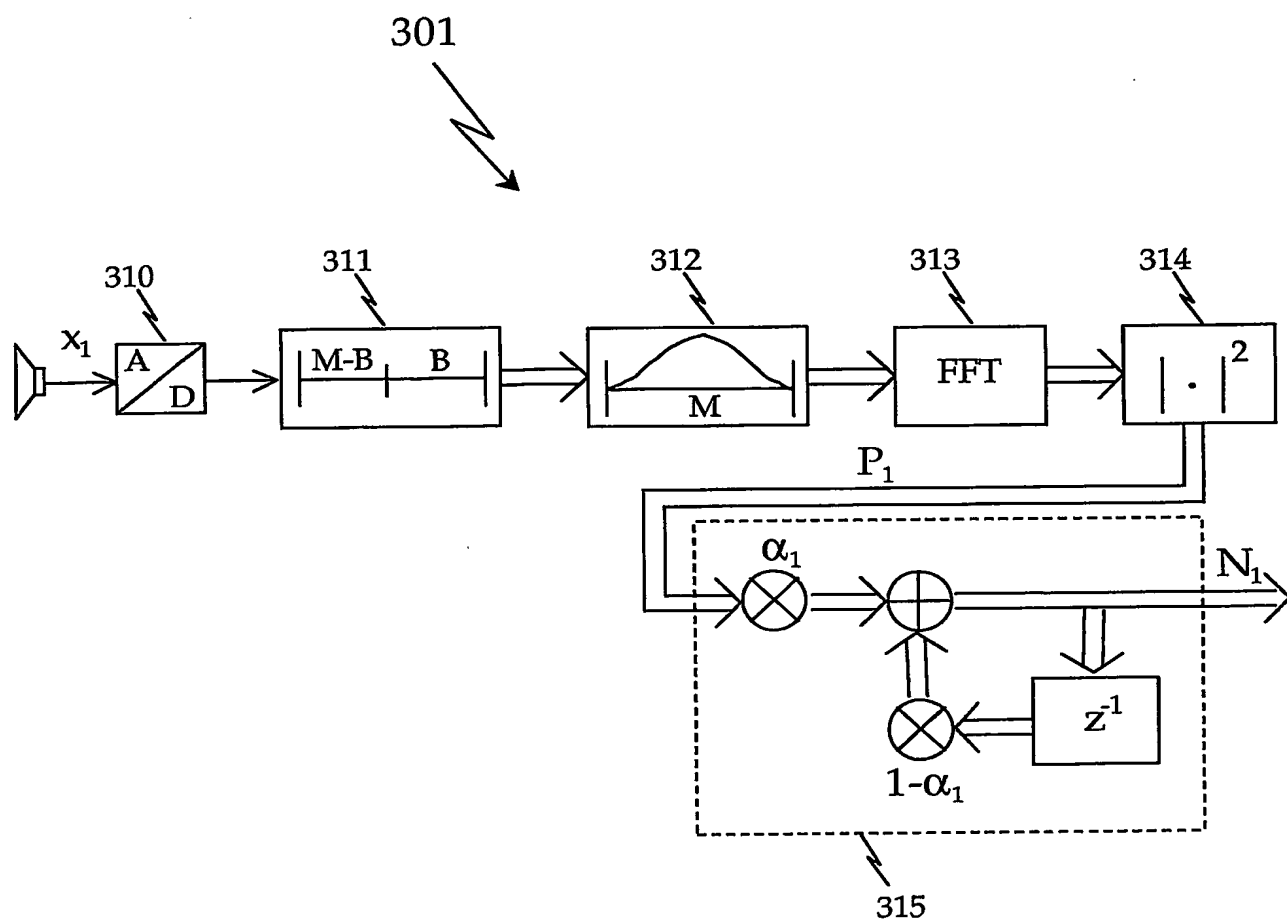
## ZUSAMMENFASSUNG

### Verfahren zur Reduktion eines Störgeräuschsignalsanteils in einem Mikrofonsignal

- Die Erfindung offenbart ein Verfahren zur Reduktion eines Störgeräuschsignalanteils in einem Mikrofonsignal, welches auf der Schätzung des Störgeräuschsignalanteils aus einem nahezu
- 5   reinen Störgeräuschsignal beruht und welches ohne zusätzliche Mikrofone auskommt.
- Wesentlich an dem erfindungsgemäßen Verfahren ist, dass die Aufnahme des Signals, welches als Basis für die Schätzung des Störgeräuschsignalanteils in dem interessierenden Mikrofonsignal dient mittels eines oder mehrerer invers betriebener Lautsprecher erfolgt.
- Insbesondere in Situationen, in denen bereits einer oder mehrere Lautsprecher als
- 10   Komponente eines Audio Systems vorhanden sind, entfällt die Notwendigkeit der Installation weiterer Mikrofone. Eine solche Situation ist beispielsweise in jedem Kraftfahrzeug, welches mit einer Audio Anlage ausgestattet ist, gegeben.



**Fig. 1**

**Fig. 2**

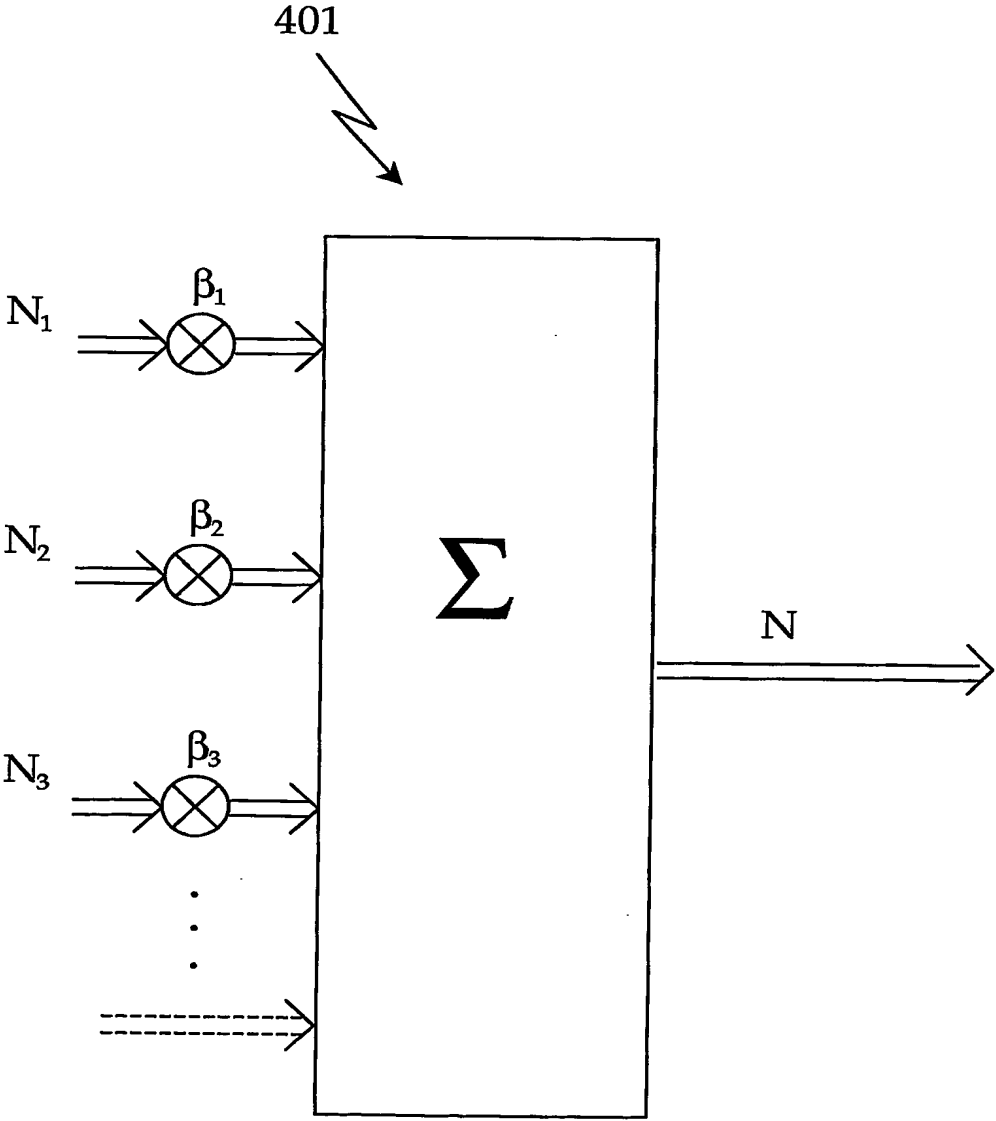


Fig. 3

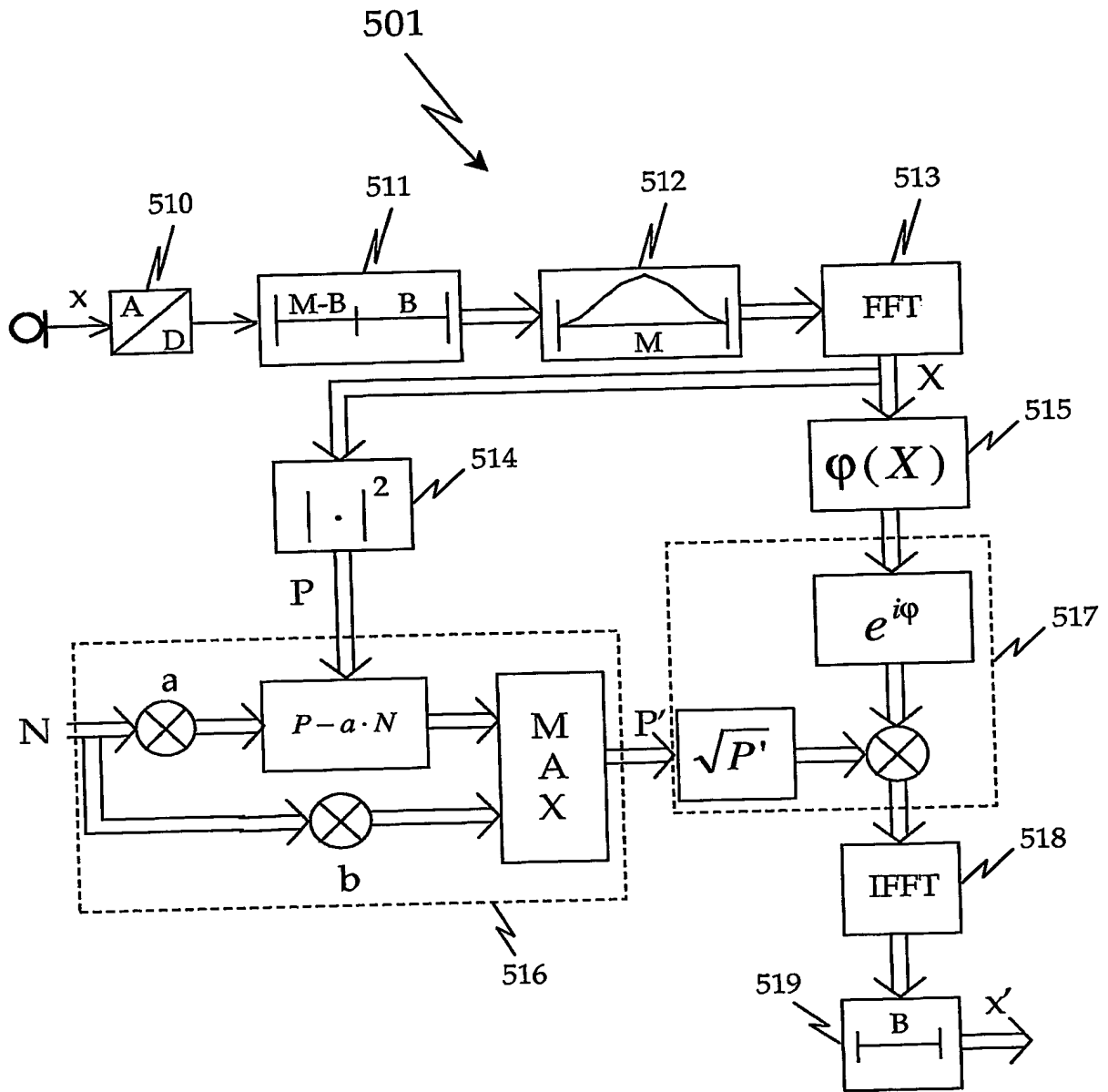
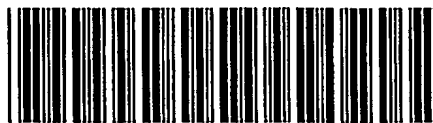


Fig. 4

**PCT/IB2004/001025**



This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images  
problems checked, please do not report the  
problems to the IFW Image Problem Mailbox**